IDU Workshop Presentation 2002

Application of Generalized Tikhonov Regularization to Earth Ecosystem Data-Model Fusion

A. J. Meade
Department of Mechanical Engineering
and Materials Science
Rice University
URL: www.ruf.rice.edu/~meade

Introduction

We are developing a method using Generalized Tikhonov Regularization (GTR) and Support Vector Machines (SVM) for seamlessly fusing a-priori models and multi-dimensional observable data from Earth Science problems.

Introduction

The accurate and efficient formation of multi-dimensional functions from observable data is of considerable importance in engineering and has a number of immediate applications which include:

- **Data acquisition**
- **Classification**
- **Controls**
- **Image recognition**

Introduction

This is the promise of artificial neural networks (ANNs) as applied to engineering

$$f_a(\underline{x}) \cong \sum d_j y_j \left(\sum b_k y_k(\underline{x})\right)$$

However, the approximation of a function $F(\underline{x})$ by a general weighted series of bases $f_a(\underline{x})$ is an *ill-posed* problem.

Introduction

Say our desired approximation minimizes the following objective function,

$$\frac{1}{2}\sum_{k}^{S}\left(F_e(\underline{x}_k) - f_a(\underline{x}_k)\right)^2 = \left\|\underline{F}_e - \underline{f}_a\right\|_2^2$$

where $F_e(\underline{x}) = F(\underline{x}) + m(\underline{x})$

Specifically, there are an infinite combination of parameters in $f_a(\underline{x})$ **that can work and minute perturbations to the input or output,** m**, can result in large changes in the approximation.**

Introduction

This ill-posed problem explains why ANNs can have trouble converging even after one has settled on

- **The architecture**
- **The number of hidden layers**
- **The type of transfer function**
- **The number of nodes**

Approach / GTR

A combined remedy is to constrain the parameters and transfer functions to form well-behaved basis functions, e.g.,

$$f_a(\underline{x}) \cong \sum c_i \Phi_i (a_i \underline{x} - q_i)$$

and to apply *regularization*. The most common and well-known form is that of Tikhonov regularization,

$$\underline{f_a} = \operatorname{argmin} \left\{ \frac{1}{2} \left\| \underline{F_e} - \underline{f_a} \right\|_2^2 + h^2 \left\| \mathbf{A}\, \underline{f_a} - \mathbf{A}\, \underline{g} \right\|_2^2 \right\}$$

where $\underline{g}$ is the vector form of the a-priori function, n is the regularization parameter and $\mathbf{A}$ is either the identity matrix or a discrete approximation of a linear derivative operator.

Approach / GTR

In our Generalized Tikhonov regularization (GTR) we are not limited to the L2 norm and we can utilize nonlinear differential operators $L[\ \ ]$ in energy form,

$$\min \left\{ \frac{1}{2} \left\| \underline{F_e} - \underline{f_a} \right\|_Y^2 + \frac{1}{2}\, \boldsymbol{l} \Lambda\left(f_a, g\right) \right\}$$

In our approach to GTR, we keep the L2 norm and directly solve the optimization problem for $f_a(\underline{x})$, i.e.

$$f_a(\underline{x}) = g(\underline{x}) + \sum_{\mathrm{i}} G\left(\underline{x}, \underline{x}_i\right) c_i \quad \text{where} \quad \underline{c} = \left[\mathbf{G} + \boldsymbol{l}\mathbf{I}\right]^{-1} \left(\underline{F_{\mathrm{e}}} - \underline{g}\right)$$

and where $G\left(\underline{x}, \underline{x}_i\right)$ is the Green's function (GF) for the differential operator $L[\ \ ]$.

Approach / GTR

The positive scalar l **is such that** $\left\| \underline{F}_e - \underline{f}_a \right\|_Y^2 \leq r$**, which satisfies our original GTR objective function.**

However, the true optimal l **is**

$$l^* = \frac{\left\| \mathbf{G}\, \underline{m} \right\|_2}{\left\| \underline{F} - \underline{g} \right\|_2} \leq \frac{\left\| \mathbf{G} \right\|_2 \left\| \underline{m} \right\|_2}{\left\| \underline{F} - \underline{g} \right\|_2}$$

and since $\mathbf{G}$ **is positive definite then as** $g\left(\underline{x}\right) \longrightarrow F\left(\underline{x}\right)$, $l^* \longrightarrow \infty$ **and as** $m\left(\underline{x}\right) \longrightarrow 0$, $l^* \longrightarrow 0$.

Approach / GTR

To summarize, with the GTR formulation:

$$f_a(\underline{x}) = g(\underline{x}) + \sum_{\mathrm{i}} G\left(\underline{x}, \underline{x}_i\right) c_i \quad \text{where} \quad \underline{c} = \left[\mathbf{G} + l\mathbf{I}\right]^{-1}\left(\underline{F}_{\mathrm{e}} - \underline{g}\right)$$

- $g\left(\underline{x}\right)$ **is the a-priori information and can be which can be a physics based numerical model, analytical solution, statistical correlation, other empirical data, or even other intelligent system models (ANNs, Fuzzy-Neural Networks, etc.).**

- $G\left(\underline{x}, \underline{x}_i\right)$ **is the Green's function for the differential operator** $L[\ \ ]$.

- l **can range from** 0 to ∞ **and is dependent on** $\mathbf{G}$, $m\left(\underline{x}\right)$, **and** $\left(F\left(\underline{x}\right) - g\left(\underline{x}\right)\right)$.

Approach / GTR

In our application of GTR to ES and engineering, we make use of a variant of the SVM to finally form well-behaved bases. This SVM variant minimizes the length of the vector $\underline{c}$ and implicitly solves for l^* by satisfying the user criteria,

$$\left\| \underline{F_e} - \underline{f_a} \right\|_Y^2 \leq \boldsymbol{r}$$

The SVM operates from

$$\min \left\{ \mathbf{G}\,\underline{c} - \left(\underline{F_{\mathrm{e}}} - \underline{g} \right) \right\} \ \text{s.t.} \ \mathrm{N} \ < \mathrm{S}$$

Approach / Our SVM Algorithm

Our SVM solution for $\underline{c}$ borrows from previous work on "mesh-free" finite elements.

Minimize $\left\langle \left(\underline{R}_{k-1} - a_k \underline{G}_k\right), \left(\underline{R}_{k-1} - a_k \underline{G}_k\right) \right\rangle$ **for basis parameters.**

Solve $c_k = \dfrac{\left\langle \underline{G}_k, \underline{R}_{k-1} \right\rangle}{\left\langle \underline{G}_k, \underline{G}_k \right\rangle}$ **and update** $\underline{R}_k = \underline{R}_{k-1} - c_k \underline{G}_k$ **, where**

$a_k = \left\| \underline{R}_{k-1} \right\|_\infty$ **and** $\left[\underline{G}_1, \ldots \underline{G}_k, \ldots, \underline{G}_N \right]^{\mathrm{T}} = \mathbf{G}$ **.**

Approach / Our SVM Algorithm

This relatively painless approach

- **Minimizes the length of** $\underline{c}$ **like conventional SVM training.**
- **Requires only one user-determined parameter** $t \leq \left\| \underline{F_e} - \underline{f_a} \right\|_\infty$ **.**
- **Avoids matrix manipulation.**
- **Requires only the storage of the sample vectors** $\underline{x}_i$ **and residual vector** $\underline{R}_{k-1}$ **.**
- **Input vectors don't have to be normalized.**

Approach / GTR with SVM

Levels of GTR applied to ES and engineering:

$$\min \left\{ \mathbf{G}\underline{c} - \left(\underline{F_e} - \underline{g} \right) \right\} \text{ s.t. } \mathrm{N} < \mathrm{S}$$

1. **'I haven't a clue.' Set $g\left(\underline{x}\right)$ to zero and $G\left(\underline{x}, \underline{x}_i\right)$ to an infinitely differentiable function.**

2. **'I think I know at least how F_e behaves wrt to at least one of the variables', e.g. time and the diffusion equation. Set $g\left(\underline{x}\right) = 0$ and $G\left(\underline{x}, \underline{x}_i\right)$ to the GF of the low fidelity solution.**

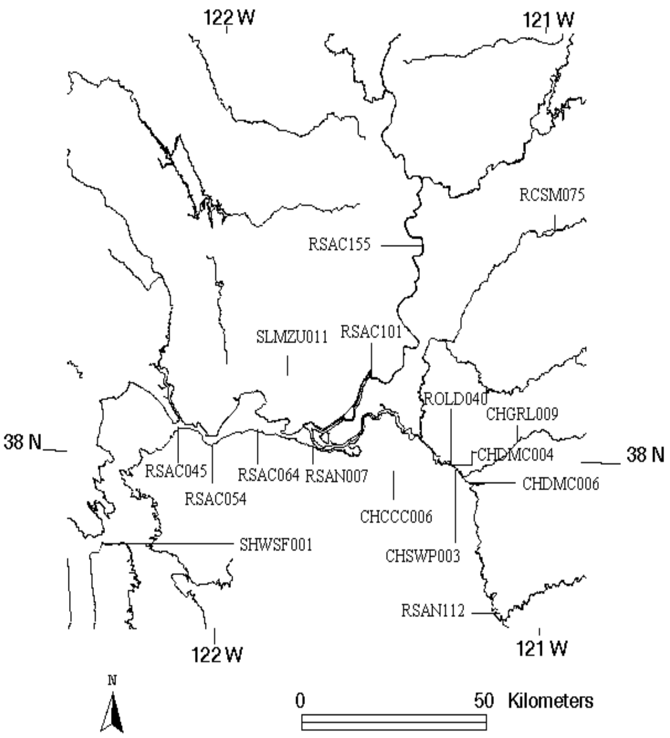
3. **'I know what's going on but need to tune the model.' Use $g\left(\underline{x}\right)$ and the GF for the model.**

Applications of GTR with SVM

- Regression / Identification in an Earth Science Problem (Level 2)
- Regression / Identification in a Rotorcraft Health Monitoring System (Level 2)
- Prediction in a Rotorcraft Health Monitoring System (Level 2)
- Classification / Identification of Naval Rotorcraft Launch and Recovery (Level 1)
- Classification / Identification in a Transonic Cavity Flow Experiment (Level 1)

Regression / Identification in an Earth Science Problem

- Inputs
  - Δ Salinity – 1 location
  - Δ Stage – 8 locations
  - Δ Flow – 6 locations
  - Δ Bias – Moon phase illumination
- Outputs
  - Δ Salinity – 1 location at next time step
- Data set (1 Hr. interval)
  - Δ Training data set – 1995-1997
  - Δ Prediction data set – 1997-1998



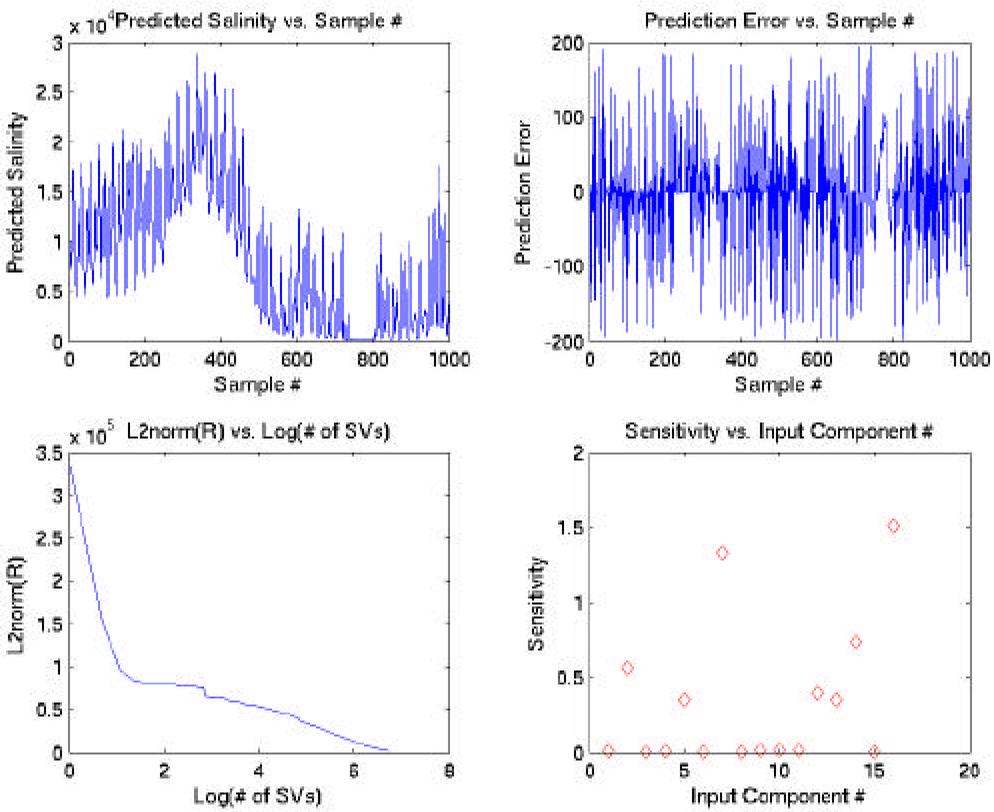
Regression / Identification in an Earth Science Problem

Our SVM model was trained using:

- **1000 sample points**
- **16 inputs**
- **1 output (water salinity)**
- t **= 200**

Can we use an SVM model with Level 2 GTR/SVM to identify dependence in this spatial time-series problem?

Regression / Identification in an Earth Science Problem



SVM model of the salinity measured at the testing station. 1000 sample points

#SV = 839

SVM Input Components

1. chgr1009-stage
2. flow-cfs-rsan112 (4)
3. stage-ft-rsan112
4. stage-feet-rsan007
5. rcsm075-flow-cfs (6)
6. rcsm075-stage-feet
7. flow-cfs-rsac155 (2)
8. stage-feet-rsac155
9. stage-feet-rsac101
10. slmzu011-stage
11. chc006
12. shwf001-stage (5)
13. chdmc00 (7)
14. chswp003 (3)
15. Moon Illumination
16. EC-rsac054 (1)

Conclusions / Future Work

The approach shows some promise. Further investigation of the method and the applications are required:

- **Investigate the physical reasoning behind the sensitivity plots from the applications presented.**
- **Make the techniques more efficient for large numbers of input samples.**
- **Use more sophisticated optimization routines.**
- **Investigate other types of bases.**